

Abstract

Previous research suggests impaired metacognitive monitoring and mathematics under-achievement in autism spectrum disorder (ASD). Within educational settings, metacognitive monitoring is supported through the provision of feedback, (e.g., with goal reminders and by explicitly correcting errors). Given the strength of the relationship between metacognition, learning and educational attainment, the current research tested new computer-based metacognitive support (the 'Maths Challenge') for mathematics learners with ASD within the context of their classroom. The Maths Challenge required learners to engage in metacognitive monitoring before and after answering each question (e.g., intentions and judgments of accuracy), and negotiate with the system the level of difficulty. Forty secondary school children with ASD and 95 typically developing learners completed the Maths Challenge in either a Feedback condition, with metacognitive monitoring support regarding the accuracy of their answers, goal reminders and strategy support, or with no feedback. Contrary to previous findings ASD learners showed an undiminished ability to detect errors. They did, however, demonstrate reduced cohesion between their pre- and post-test intentions. Support from the Feedback condition significantly improved task performance for both groups. Findings highlight important implications for educational interventions regarding the provision of metacognitive support for ASD learners to ameliorate under-performance in mathematics within the classroom.

Supporting metacognitive monitoring in mathematics learning for young people with autism spectrum disorder: A classroom-based study

Autism spectrum disorder (ASD) is characterised by impairments in social communication and behavioural flexibility and is estimated to affect around 1% of the population (American Psychiatric Association, 2013; Baird et al., 2006). These impairments may impact upon learning, and recent reports indicate greater gaps in attainment for pupils with ASD compared to their typically developing (TD) peers (Keen, Webster & Ridley, 2016; Wilkinson & Twist, 2010), with many demonstrating a specific deficit in mathematics (e.g., Chiang & Lin, 2006; Mayes & Calhoun, 2006). Despite reports of the existence of a small sub-group of mathematically gifted people with ASD (Aagten-Murphy et al., 2015; Chiang & Lin, 2006; Iuculano et al., 2014; Jones et al., 2009; Mayes & Calhoun, 2003; 2006), on average, mathematics ability is substantially lower among people with ASD than would be expected on the basis of IQ (Aagten-Murphy et al., 2015; Chiang & Lin, 2007; Estes, Rivera, Bryan, Cali & Dawson, 2011; Griswold et al., 2002; Mayes & Calhoun, 2003; 2006; see also Jones et al., 2009). Educational underachievement in mathematics contributes directly to less-than-optimal economic outcomes (e.g., low levels of employment) and life chances among people with ASD (Estes et al., 2011).

One of the most effective and cost efficient educational interventions is to support metacognition (Higgins et al., 2013). Metacognition can be defined as the ability to reflect upon, understand and control one's learning, or 'thinking about one's thinking' (Schraw & Dennison, 1994). It is well established that metacognition facilitates self-regulation of behaviour and learning: monitoring when or where mistakes are made means that learning strategies can be modified accordingly (e.g., revising until you are confident you know the topic). Within an educational context, research has highlighted that metacognition predicts academic achievement more powerfully than intellectual abilities (Hartwig, Was, Isaacson & Dunlosky, 2012; Thiede, 1999; Veenman, Kok & Blöte, 2005) and there is extensive evidence that developing metacognition is an effective intervention in school children within and below the 'normal' range of ability (Dunlosky, Kubat-Silman & Hertzog, 2003; Iuculano et al., 2014; Maxwell & Grenier, 2014; Schneider & Artelt, 2010; van der Stel & Veenman, 2010; see also Roebbers, Cimeli, Röthlisberger & Neuenschwander, 2012; Roebbers, Krebs & Roderer, 2014). Accordingly, teaching approaches that encourage learners to monitor, evaluate and strategize their learning are now widely recommended (Higgins et al., 2013; Special Education Support Service, 2009).

This is pertinent because research indicates that autistic individuals have a metacognitive deficit (Bebko & Ricciuti, 2000; Brosnan et al., 2015; Farrant, Blades & Boucher, 1999; Farrant,

Boucher & Blades, 1999; Grainger, Williams & Lind, 2014, 2016; McMahan et al., 2016; Vlamings et al., 2008; Wojcik, Moulin & Souchay, 2013). For example, children with ASD have specific metacognitive difficulties in identifying when they have made mistakes – they report more confidence in the accuracy of their answers, even when they are incorrect (Brosnan et al., 2015). This is striking because metacognitive monitoring relates to self-regulation of learning (e.g. Isaacson & Fujita, 2006) and a deficit in metacognitive monitoring could explain why children with ASD have difficulties with self-regulation of learning-related behaviour; not knowing what you know and what you do not know, or not knowing if a mistake has been made severely impairs options for responding and adapting behaviour accordingly.

Brosnan et al. (2015) found that, as well as being more likely to think an erroneous answer was correct, when told they had made an error learners with ASD were significantly more likely to report that they had *meant* to make that error. This impaired intention monitoring in ASD has also been found outside of the educational context in other types of tasks such as knee reflex reactions (Williams & Happé, 2010) and matching goals with desirable but unintended outcomes (Phillips, Baron-Cohen & Rutter, 1998). Although Brosnan et al. suggested their finding that ASD learners reported they meant to get a question wrong was an intention monitoring error, there was no assessment of intention prior to answering the mathematics questions. It is possible therefore that some learners did in fact intend to get the answer wrong (thus when they report this after the event, this was not an intention monitoring error).

Research with TD children shows that learning can be improved with the provision of immediate formative feedback that clarifies goals and minimises uncertainty in relation to how well learners are performing on a task, and what strategies they could take to achieve their goals (Azevedo & Bernard, 1995; Bangert-Drowns, Kulik, Kulik & Morgan, 1991; Shute, 2008). Metacognitive self-regulation training in planning, monitoring and control is also successful in improving mathematics learning with both TD pupils (Desoete & Veenman, 2006) as well as those with learning difficulties (Cornoldi, Lucangeli, Caponi, Falco, Focchiatti & Todeschini, 1995; Kroesbergen & Van Luit, 2003). Such training has been shown to be particularly successful when combined with cognitive and motivational strategies (Dignath et al., 2008) and when implemented via the use of a computer in a cognitive apprenticeship-style learning environment (Teong, 2003; Xin & Jitendra, 1999; Zimmerman & Tsikalas, 2005). Although there is some evidence that interventions such as reinforcements for successful task completions can support ASD learning (e.g., Adcock & Cuvo, 2009; Charlop, Kurtz & Milstein, 1992; Chong & Carr, 2005; Dunlap & Koegel, 1980), no research has specifically targeted support for metacognition and self-regulated learning in ASD. The

provision of appropriate metacognitive support for learners with ASD needs to be informed by where their difficulties lie specifically with respect to metacognitive monitoring.

The current research therefore examined: 1) the nature of the metacognitive difficulties in ASD with respect to monitoring the accuracy of answers, intentions, and regulating learning strategy accordingly; and 2) the use of feedback as metacognitive monitoring support in mathematics learners with ASD. Intentions to obtain a correct answer were assessed before as well as after each question was attempted. In addition, learners were able to adjust the level of difficulty at various points throughout the program. Mathematics performance was explored under two conditions – one condition provided metacognitive monitoring support through feedback and one did not. We predicted that the provision of metacognitive monitoring support would enhance strategy regulation (appropriately adjusting the level of difficulty), and thus mathematics performance (to obtain more points) of learners with ASD who have a deficit in metacognitive monitoring.

Different types of learning environments in and outside classrooms impact upon self-regulation in learners and research should be sensitive to this context (Boekaerts & Corno, 2005). Despite the many benefits of tightly-controlled laboratory-based research, such a paradigm can impact upon the participant's perceived value of the task, and consequently the variables under investigation. Within the classroom, intrinsic value is strongly related to use of cognitive strategies and self-regulation by typically developing learners. For example, Pintrich and de Groot (1990) report that learners who are cognitively engaged and self-regulating are those who are interested in and value their classroom academic work. It was therefore felt crucial to examine metacognition and self-regulation within the context of the classroom, specifically a typical mathematics lesson, while acknowledging that this will limit the extent to which ASD and TD learners can be considered matched. Whether and how to match is an ongoing issue in ASD research (e.g., Barbeau, Soulières, Dawson, Zeffiro & Mottron, 2013; Jarrold & Brock, 2004). Importantly, the literature above suggests that those with ASD will have impaired mathematical ability and metacognition and therefore matching on these (for example with a learning disabilities group - who may also have impaired language) is problematic. Matching mathematics ability with younger learners can also be problematic as the metacognitive factors under investigation do not predict ability until around 11 years of age (Roebers et al., 2014) and learners with ASD can make types of mathematics errors that are not evidenced in TD learners (and vice versa; Brosnan et al., 2015). In an attempt to address these issues, the present study's methodology presented mathematics questions at the appropriate level for each student individually in the context of their classroom-based mathematics lesson through a computer-based 'Maths Challenge' (described below).

Method

Participants and design

Forty secondary school children (30 male, 10 female) who had received a formal diagnosis of autism or Asperger syndrome by a qualified clinician according to DSM (American Psychiatric Association, 2000, 2013) or ICD criteria (World Health Organisation, 1993) were recruited from schools or units specifically for children with ASD. Secondary School in England covers ages 11-16 where the curriculum is divided between Key Stage 3 (ages 11-14) and Key Stage 4 (ages 15-16). Under-achieving children may still be working at Key Stage 2 (for ages 7-11) or even Key Stage 1 (for ages 5-7)¹. The mean age of the ASD group was 13.33 years ($SD = 1.25$, range = 11-16 years), and the majority were working below the expected level, at Key Stages 1 ($n = 10$; mean age = 12.70 years) or 2 ($n = 20$; mean age = 13.35 years)², however 4 ASD pupils were working at Key Stage 3 (mean age = 13.50 years) and 6 were working at Key Stage 4 (mean age = 14.17 years). All children with ASD were educated within specialist provision classrooms housed within mainstream schools. Comparison participants comprised 95 secondary school pupils (58 male, 37 female) from mainstream schools, with a mean age of 13.40 years ($SD = 1.15$, range = 11-15 years), and all were working at the appropriate level, at Key Stages 3 ($n = 64$, mean age = 12.88 years) and 4 ($n = 31$, mean age = 14.48 years). Ethical approval was prospectively obtained from the University of Bath research ethics committee.

It is important to note that the groups did not significantly differ on age, $t(133) = .34$, $p = .74$, $d = .06$, or proportion of males to females, $\chi^2 = 2.41$ (1, $N=135$), $p = .12$, $\phi = .13$, and the ASD group was under-achieving in mathematics as has widely been reported in the literature (e.g., Aagten-Murphy et al., 2015; Chiang & Lin, 2007; Estes et al., 2011; Griswold et al., 2002). The samples were therefore reflective of the populations from which they were drawn. The focus of the present study was not on absolute level of mathematical ability – as the computer program was flexible to start at the appropriate level for each student (Key Stage 1 to 4) – but on the influence of feedback on metacognition (see Discussion). The study took place within the context of the classroom and assessments of mathematical ability were from the classroom teacher and starting level mathematics performance on the current task, as formal assessments of mathematical ability were not possible within this context.

¹ <https://www.gov.uk/government/publications/national-curriculum-in-england-mathematics-programmes-of-study>

A 2 (Group: ASD vs. TD) x 2 (Metacognitive support: Feedback vs. No feedback) between participants design was used, whereby participants were randomly assigned to either a 'Feedback' or a 'No Feedback' condition (see below for more details), with the constraint that there were approximately the same number of participants in each group x condition cell. The final sample comprised 40 ASD participants (21 feedback; 19 no feedback) and 95 TD participants (46 feedback; 49 no feedback). There was an even distribution in the ratio of males to females in each feedback condition for both the ASD ($p = .57$, $\phi = .03$, Fishers exact test) and TD group ($\chi^2 (1, N = 95) = .21$, $p = .65$, $\phi = .05$), and age was evenly distributed between conditions for both ASD, $F(1, 38) = .04$, $p = .84$, $\eta p^2 = .001$, and TD groups, $F(1, 93) = .18$, $p = .67$, $\eta p^2 = .002$. There was an approximately even distribution of participants working at the different Key Stages between conditions for both the ASD ($p = .51$, $\phi = .26$, Fishers exact test) and the TD group ($\chi^2 (1, N = 95) < .001$, $p > .99$, $\phi = < .001$) (Table 1). Finally, while starting level mathematics performance, in terms of number of correct answers provided for the first block of questions (that is, independent from metacognitive ability and feedback condition), was significantly lower among ASD participants compared to the TD group, $F(1, 131) = 9.90$, $p = .002$, $\eta p^2 = .07$, it did not differ between feedback conditions, $F(1, 131) = 1.02$, $p = .32$, $\eta p^2 = .01$. A lack of Group x Condition interaction indicated that starting level mathematics did not significantly differ between feedback conditions for either group, $F(1, 131) = 1.17$, $p = .28$, $\eta p^2 = .01$. Thus, we were confident that age, sex and mathematical ability did not significantly differ between the Feedback and No Feedback conditions.

Table 1. Distribution of participants, mean ages and ratio of males to females working between the different Key Stages in each condition

		ASD				TD			
		Key Stage 1	Key Stage 2	Key Stage 3	Key Stage 4	Key Stage 1	Key Stage 2	Key Stage 3	Key Stage 4
Feedback	% (n)	15% (6)	30% (12)	2.5% (1)	5% (2)	0	0	33% (31)	16% (15)
	M age (SD)	12.83 (.41)	13.25 (1.22)	14.00 (-)	14.50 (.71)			12.84 (.93)	14.40 (.63)
	N male; female	5; 1	8; 4	1; 0	2; 0			17; 14	10; 5
No Feedback	% (n)	10% (4)	20% (8)	7.5% (3)	10% (4)	0	0	35% (33)	17% (16)
	M age (SD)	12.50 (1.29)	13.50 (1.93)	13.33 (1.15)	14.00 (.00)			12.91 (1.01)	14.56 (.63)
	N male; female	3; 1	5; 3	2; 1	4; 0			21; 12	10; 6

Materials and procedure

The Maths Challenge. We developed the “Maths Challenge” computer program in Real Studio (Xojo, 2011), whereby, to maximise points won, participants needed to monitor their performance and adapt their strategy accordingly. The program comprised seven levels of difficulty and questions were more difficult the higher the level. Each question answered correctly was worth points commensurate to that level (e.g., 3 points per question answered correctly on Level 3), while errors were worth 0 points. Participants answered four blocks of three questions. In order to provide participants the opportunity to move up or down in level of difficulty from the outset (after completing the first block), all participants started the game at Level 4. After completing each block, participants decided themselves whether to stay at the same level (with the same points available per question), move up a level (with more points available per question) or move down a level for the next block (with fewer points available per question) and so forth until they had completed all four blocks. Thus, notwithstanding mathematics ability, an individual’s maximum points potential was determined by their metacognitive strategy regulation (e.g., if all questions on Level 4 were answered incorrectly the most rewarding strategy would be to move down to Level 3 for the next block), as well as their ability to judge when errors are made.

Four versions of the Maths Challenge were developed to accommodate pupils at varying levels of mathematical ability in terms of the English Key Stage (KS) level they were currently working at. Specifically, Version 1 comprised questions from KS1; Version 2 questions from KS2; Version 3 questions from KS3; and Version 4 questions from KS4. Within each version, Levels 1 to 7 comprised questions from lower to higher KS sub-levels within that respective Key Stage. Mathematics questions were selected from UK National Curriculum past test papers and exam revision workbooks on the basis that they could be answered mentally, without the need for pen, paper or calculator. The program began with general instructions explaining the structure, aims and points allocations. Each mathematics question was preceded by a 5-point pre-test intention measure (“how hard are you going to try to get the next question right”), and was followed by a 5-point post-test metacognitive monitoring confidence judgment (“do you think you got that question right or wrong”), and finally a 3-point post-test intention measure (“did you mean to get that question right or wrong”) (Figure 1). For the purposes of the analyses, the 5-point pre-test measure was collapsed into 3 points (1 = try to get wrong; 2 = neutral; 3 = try to get right). A prototype of the program was piloted with both ASD and TD children and received positive evaluations, indicating that children were motivated to perform well.



Figure 1. Order of pre- and post-test intention and metacognitive judgment measures for each mathematics question in the Feedback condition. Note that in the No Feedback condition the correct answer (the fourth display box) was omitted.

Participants were tested in classroom groups during school lessons; however, each pupil completed the program individually at their own computer. In the Feedback condition, participants received feedback after each question regarding whether they had answered correctly or not, and how many points they had won for that question. This was displayed in the text and also graphically with gold coins denoting the points won for that question, as well as a running total of points won thus far displayed in a box on the top left corner of the screen. After completing each level, Feedback participants also received a summary regarding the number of questions they had answered correctly and the number of points won, which was accompanied by a goal reminder that their task was to finish the program with as many points as possible. After completing the three questions in each level, all participants were asked whether they wanted to move up, down, or stay on the same level. In the feedback condition, this was accompanied by strategy reminders; for example that choosing to go down a level would mean easier questions, but fewer points available for each question. To examine metacognitive monitoring (of mathematics accuracy) and strategy regulation (level decisions that optimise correct answers and points won) in the absence of external

support, the No Feedback condition provided no indication of whether questions had been answered correctly or not, and information regarding points won and goal/strategy reminders were omitted from the display (Figure 2).

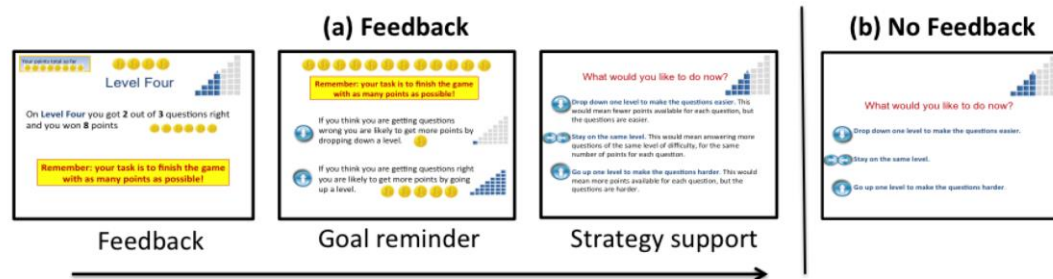


Figure 2. Provision of feedback, goal reminders and metacognitive strategy support prior to level decisions in the Feedback condition (panel A). In the No Feedback condition participants were simply asked to make a level decision (panel B).

Finally, participants completed a brief metacognitive questionnaire, comprising four questions (each answered on a 4-point scale) reflecting on their metacognitive performance, specifically: awareness of performance (from 'never or rarely aware' to 'always aware'); points plan (from 'no plan' to 'get as many points as possible'); level strategy (from 'stayed on the same level regardless of difficulty' to 'if easy then up; if difficult then down a level'); checking of answers (from 'never or rarely checked' to 'always checked answers').

Results

Analysis overview

We report the results of analyses conducted to test the nature of the metacognitive difficulties in ASD (and whether experimental task performance differed between groups and feedback conditions) with respect to the following metacognitive variables: Judgments of confidence for correct and incorrect answers; pre- and post-question intention monitoring; and self-reported metacognition ratings. Finally, we report the effect of feedback in terms of the number of correct answers in the first and final blocks. Table 2 provides a summary of the means and SDs of each of these variables between groups and conditions.

Table 2. Mean scores across all task variables (standard deviations are in parentheses)

	ASD		TD	
	No Feedback	Feedback	No Feedback	Feedback
Judgements of confidence in answers (range = 0-1)	.69 (.22)	.76 (.25)	.70 (.17)	.67 (.25)
Mean pre-test question intentions (range = 1-3)	2.88 (.35)	3.00 (.02)	2.90 (.30)	2.74 (.54)
Mean post-test intentions (range = 1-3)	2.51 (.66)	2.49 (.59)	2.71 (.52)	2.36 (.70)
Self-reported metacognition score (range = 0-12)	4.53 (3.49)	3.38 (3.17)	4.55 (2.37)	4.13 (3.57)
Starting level Mathematics performance (number of points won on first block) (range = 0-12)	5.26 (4.43)	6.86 (4.76)	8.49 (3.80)	8.43 (3.79)
Final level Mathematics performance (number of points won on final block (range = 0-21)	3.21 (5.13)	7.67 (6.69)	6.29 (5.73)	7.30 (5.27)

Judgments of confidence

Participants' ordinal metacognitive judgment ratings of their accuracy to each question on a 5-point confidence scale ('I'm sure I got it right'; 'I think I got it right'; 'I don't know'; 'I think I got it wrong'; 'I'm sure I got it wrong') were translated question-by-question into interval data (1, 0.75, 0.5, 0.25, 0) for the analyses.

To determine whether participants assigned higher confidence to correct compared to incorrect answers, a 2 (Group) x 2 (Condition: Feedback vs. No Feedback) x 2 (Answer: correct vs.

incorrect) mixed-ANOVA was conducted, where Answer was the within-participants factor³. There was a main effect of Group, $F(1, 120) = 4.15, p = .04, \eta p^2 = .03$, whereby the ASD group were significantly more confident generally that their answers were correct ($M = .73, SD = .23$) than were the TD group ($M = .68, SD = .21$). There was also a significant main effect of Answer, $F(1, 120) = 87.34, p < .001, \eta p^2 = .42$, whereby confidence was significantly higher for correct answers ($M = .80, SD = .18$) than for incorrect answers ($M = .56, SD = .27$). There was no effect of Condition, $F(1, 120) = .11, p = .74, \eta p^2 = .001$, Group x Condition interaction, $F(1, 120) = 2.69, p = .11, \eta p^2 = .02$, or Group x Condition x Answer interaction, $F(1, 120) = 1.60, p = .21, \eta p^2 = .01$, indicating that the provision of feedback did not affect the nature of metacognitive judgements among either group. Finally, the Group x Answer interaction was not significant, $F(1, 120) = .59, p = .45, \eta p^2 = .005$, indicating that both groups showed higher confidence for correct than incorrect answers (Figure 3).

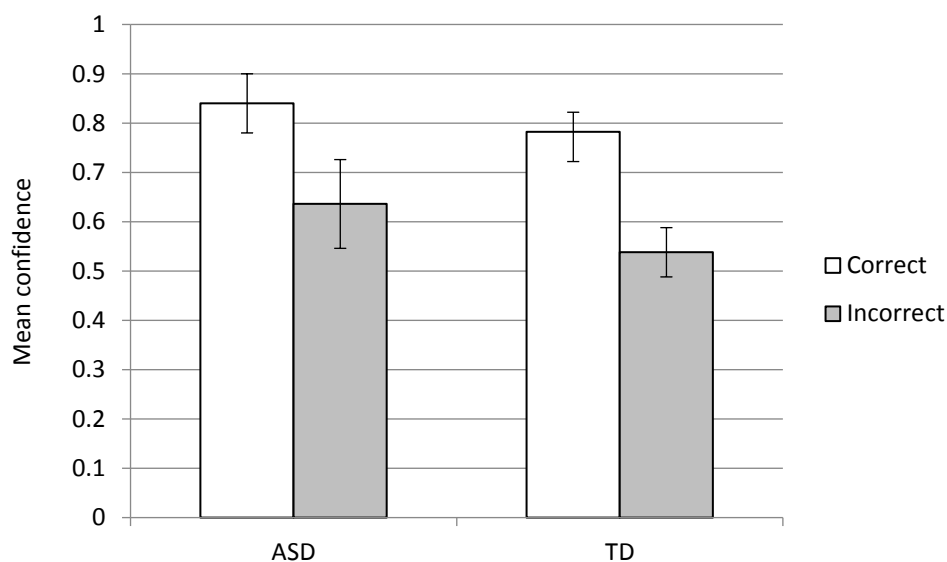


Figure 3. Mean confidence in correct and incorrect answers by ASD and TD participants (error bars indicate the 95% confidence intervals around the mean)

Intention monitoring

Pre-test intentions. A 2 (Group) x 2 (Condition) between participants ANOVA revealed no significant main effects of Group, $F(1, 131) = 1.82, p = .18, \eta p^2 = .01$, or Condition, $F(1, 131) = 1.23, p = .27, \eta p^2 = .01$. The Group x Condition interaction for pre-test intentions was also not significant, $F(1, 131) = .57, p = .45, \eta p^2 = .004$.

³ Four TD participants and two ASD participants did not produce any correct answers, and four ASD participants did not produce any incorrect answers. These participants were excluded from the analysis.

Post-test intentions. A 2 (Group) x 2 (Condition) between participants ANOVA revealed no significant main effects of Group, $F(1, 131) = .06, p = .80, \eta p^2 < .01$, or Condition, $F(1, 131) = 2.56, p = .11, \eta p^2 = .02$. The Group x Condition interaction for pre-test intentions was also not significant, $F(1, 131) = 2.06, p = .15, \eta p^2 = .02$.

Predicting post-test intentions from pre-test intentions. Pre- and post-test intentions were significantly positively correlated for both ASD, $r = .64, p < .001$, and TD groups, $r = .81, p < .001$. To examine whether the strength of the association between pre- and post-test intentions was similar for both groups, Fisher's Z transformations were conducted. These indicated a significantly weaker association between pre- and post-test intentions for the ASD group compared to the TD group, $Z_{r1-r2} = 1.89, p = .03$. Figure 4 displays average post-test intentions (from 1 = 'I meant to get it wrong' to 3 'I meant to get it right') for each pre-test intention option (from 1 = 'I'm going to try to get it wrong' to 3 = 'I'm going to try to get it right') for errors made by ASD and TD groups. The TD group showed better calibration (with scores closer to the dotted line) than the ASD group.

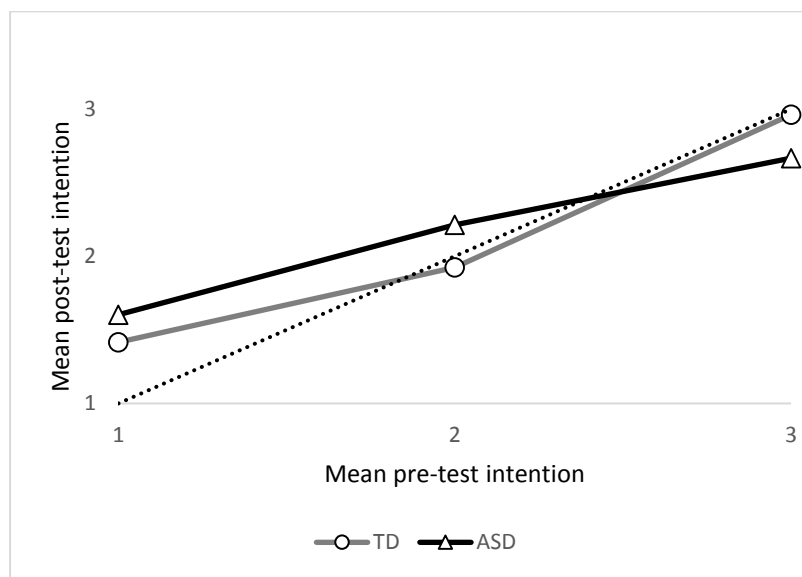


Figure 4. Pre- and post-test intentions for errors. The dotted line indicates perfect calibration.

Metacognitive questionnaire

A 2 (Group) x 2 (Condition) MANOVA was conducted with participants' four self-reported metacognitive ratings (awareness of performance; points plan; level strategy and checking of answers) as dependant variables. There was no multivariate effect of Group, $F(4, 128) = 1.63, p = .17, \eta p^2 = .05$, or Condition, $F(4, 128) = 2.13, p = .08, \eta p^2 = .06$, and the Group x Condition interaction was also not significant, $F(4, 128) = .61, p = .66, \eta p^2 = .02$.

Mathematics performance: Effect of feedback

Groups significantly differed on starting level mathematics, as described above and evidenced by significantly fewer points won by the ASD group on the first block of questions, $F(1, 131) = 9.90, p = .002, \eta p^2 = .07$, with no main effect of Condition or Group x Condition interaction at the start of the task ($ps > .28$). To examine whether feedback was effective in improving self-regulation, as measured by optimal strategy use to win as many points as possible, a 2 (Group) x 2 (Condition) between participants ANOVA was conducted for points won on the final block⁴. There was a significant effect of Condition, $F(1, 131) = 6.57, p = .01, \eta p^2 = .05$, whereby more points were won in the Feedback ($M = 7.42, SD = 5.71$) compared to the No Feedback condition ($M = 5.43, SD = 5.70$) but no significant main effect of Group, $F(1, 131) = 1.61, p = .21, \eta p^2 = .01$ or Group x Condition interaction, $F(1, 131) = 2.59, p = .11, \eta p^2 = .02$. Thus, the provision of feedback improved performance for both groups (Figure 5).

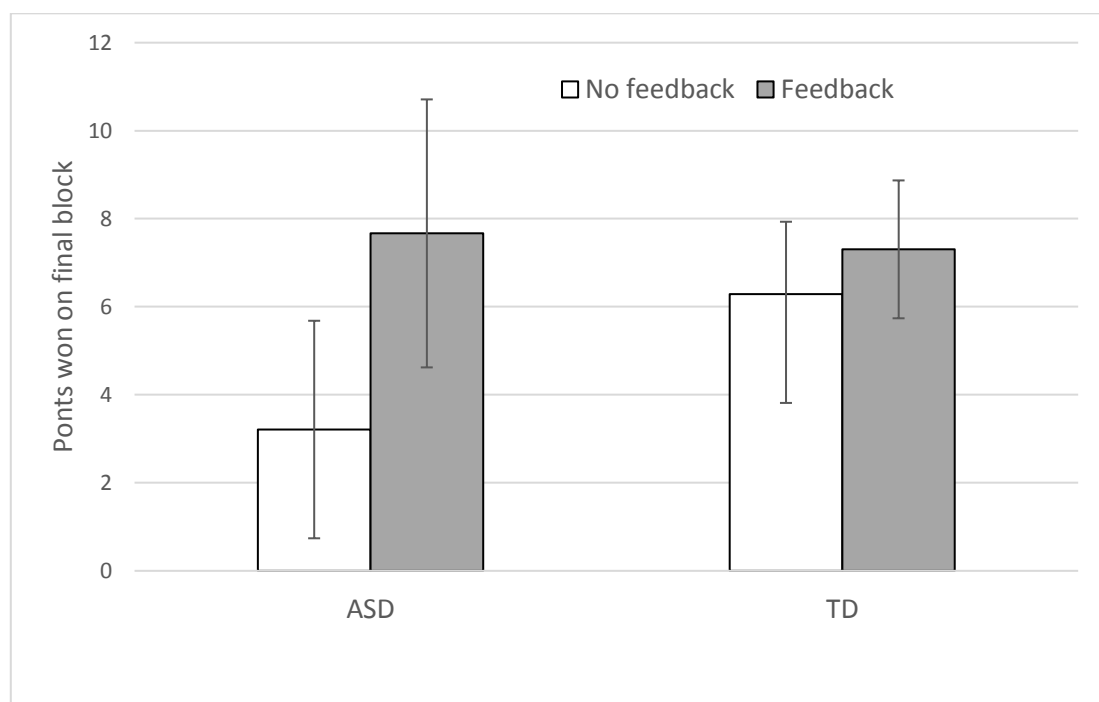


Figure 5. Number of points won on the final block in the Maths Challenge as a function of Group and Feedback condition (error bars indicate the 95% confidence intervals around the mean)

⁴ Since ASD and TD groups differed on starting level mathematics performance (and thus the number of points won at the start of the game), strategy would be expected to have its largest effect on points won on the final block.

Discussion

'Metacognition', or *awareness* of one's cognition, makes a major contribution to the learning of mathematics and is a better predictor of educational achievement (e.g., exam performance) than assessments of intelligence. Self-regulated learning requires setting one's own goals in relation to learning and ensuring that they are attained (Efklides, 2011). Effectively regulating learning, for example through study time allocation and revision methods, is dependent on the ability to monitor what and how much material is known, as well as one's goals and intentions (Metcalf, 2009; Nelson & Narens, 1990). Despite the metacognitive difficulties and educational under attainment experienced by people with ASD, no research has specifically targeted support for metacognition and self-regulated learning in ASD. The context of self-regulated learning has been argued to be crucial, requiring research to be as context-sensitive as possible (e.g., Boekaerts & Corno, 2005); thus, the aim of the current research was to test appropriate metacognitive support for ASD mathematics learners within the context of their mathematics lessons.

The present findings indicate that ASD learners can successfully distinguish correct from incorrect answers. This is contrary to previous reports that autistic children show diminished metacognitive monitoring, reporting more confidence in their answers, even when they are incorrect (e.g., Brosnan et al., 2015; Grainger et al., 2016; McMahon et al., 2016; Williams, Bergström & Grainger, 2017). Findings from the current study indicate that while individuals with ASD showed a general bias toward higher confidence, they were still able to differentiate correct from incorrect answers. A potential explanation for the current diverging findings could lie with the types of tasks used, with the present study adopting a non-socially delivered task in the classroom context with elements from serious gaming, which together may have increased ASD learners' understanding, engagement and intrinsic motivation in the task (see Kenworthy, Yerys, Anthony & Wallace, 2008; White, Burgess & Hill, 2009; Whyte, Smyth & Schef, 2015).

While monitoring the accuracy of their answers appeared undiminished, learners with ASD nevertheless showed reduced cohesion between their pre- and post-test intentions. Including a measure of pre-test intentions enabled the present study to extend previous findings by Brosnan et al. (2015) to confirm that errors made on mathematics questions were not intended, and therefore that post-hoc reporting that errors had been intended more likely reflects diminished intention

monitoring. This is important as a diminished sense of one's intentions towards a task could have a significant impact on one's capacity to benefit from the experience of attempting the task. Interestingly there were no differences in the absolute assessments of pre- and post- intentions to answer questions correctly; the group difference emerged in the relationship between pre- and post- intentions, thus indicating a more subtle impairment in monitoring one's own intentions in ASD.

Learners with ASD benefited from metacognitive support, as evidenced by their better performance in the Feedback compared to the No Feedback condition. Although the Group x Condition interaction did not reach statistical significance, this feedback support appeared to be less effective for the TD group (Figure 5); however these young people did not have a metacognitive deficit. The error bars in Figure 5 highlight the variability in performance and the need for caution in interpreting the data, but with no feedback the ASD group average around half the number of points of the TD group and slightly more points than the TD group with the provision of feedback. The provision of feedback, goal and strategy support may be most effective for learners who have difficulties in metacognitive monitoring.

Higgins et al. (2013) report that supporting metacognition is a fast, efficient and low cost method for supporting learning. Consistent with this, the present research suggests that for learners with diminished metacognitive monitoring, metacognitive support should: (i) affirm the goal (e.g., to get the answer correct); (ii) feedback immediately (e.g., whether the answer was correct/incorrect); and (iii) reflect on the goal (e.g., intention monitoring). A limitation of the present study is that it was not possible to disentangle the relative benefits of each of these components of metacognitive support provided by the feedback condition. This is especially pertinent in the context of the present findings of an intention monitoring deficit alongside undiminished accuracy monitoring. For example, it is unclear whether the provision of feedback about performance was a necessary component in supporting task performance, or whether goal reminders alone would be sufficient. Moreover, although successful performance on the Maths Challenge was dependent on strategy regulation (to win as many points as possible), the present study did not extricate where group differences lay in metacognitive strategy regulation. Further analysis of participants' self-selecting of difficulty level, for example, may reveal differences between the two groups. If learners with ASD are also poorer at judging when to best shift difficulty levels, despite when feedback is provided, implementing Dynamic Difficulty Adjustment (DDA) (Hunicke, 2005), where game difficulty is modified automatically in response to students' answers, may be beneficial. Anderson (2012) proposed developing a framework of conceptual knowledge for the teaching of fractions in a Digital Educational Game, where the game keeps track of players' knowledge and adapts the difficulty level

accordingly. Future work may investigate differences in performance between self-selecting difficulty and DDA, and whether this can further support metacognition in ASD. Finally, the level of challenge in the current study was individualised, which may be significant in the efficacy of support (e.g., not too easy/hard). Future research should tease apart these aspects in order to more specifically target where support is needed in terms of metacognitive monitoring, intentions and self-regulation respectively.

These steps to support learning would be expected to extend to all learners with deficits in metacognitive monitoring, not just those with ASD, although this needs to be assessed empirically. There may be other reasons underpinning poor mathematical performance, such as developmental dyscalculia, which is thought to be a deficit in a basic capacity for understanding numerosity (e.g., Butterworth, 2004). People with ASD may also have additional deficits in numerosity (e.g., Aagten-Murphy et al., 2015), which one would not expect to be ameliorated through interventions addressing metacognition. The ASD group were mostly working at a Key Stage level below where they were expected to be working (the TD group were not). Thus, whilst the ASD group were found to benefit from feedback, this does not necessitate that they were consequently at the same academic level as the control group. Future research could assess both mathematical skills and metacognitive skills to establish any interrelationships. We would expect that supporting metacognitive monitoring would be the first step in addressing numerosity deficits. It is worth reiterating, however, that whilst people with ASD typically show a specific impairment in mathematical ability incommensurate with their IQ, there is also a small group of mathematically gifted people with ASD (Aagten-Murphy et al., 2015; Chiang & Lin, 2006; Luculano et al., 2014; Jones et al., 2009; Mayes & Calhoun, 2003; 2006). The present finding may not extend to such a group, and it would be interesting to identify how metacognition and numerosity skills differentiated this mathematically gifted sub group from the typical ASD profile.

Metacognitive monitoring is considered essential for a sense of 'self-concept' (Roebbers et al., 2012) and for day-to-day behavioural functioning, because accurate monitoring of one's internal states facilitates the regulation of and control over those states, and over learning and behaviour (Nelson & Leonesio, 1988). We have used the term metacognitive monitoring to encompass assessments of intention before and after each task, and accuracy, as well as the capacity to control appropriate cognitive responses to these assessments. This process has been distinguished from metacognitive knowledge of cognition, which is the stored acquired knowledge of cognition (e.g. 'I am better at arithmetic than I am at spelling'; Flavell, 1979; Lockl & Schneider, 2002; Nelson & Narens, 1990; Schraw & Moshman, 1995). Despite differences between groups in intention monitoring, there were no self-reported differences on the metacognition questionnaire. If we

assume that the intervention did effectively target metacognitive monitoring, it may be that those with ASD are unable to self-report their own metacognition. This is consistent with Williams and Happé (2010) who argue that a metacognitive understanding of one's own mind may be more impaired in ASD than the metacognitive understanding of other people's minds. Indeed, Grainger et al. (2014) report that despite behaviourally demonstrating impaired metacognition (compared to controls), people with ASD self-reported higher levels of metacognition.

It is a limitation of the present study that that we did not undertake an independent assessment of autism diagnosis. Working within the schools, this was not feasible. We used a convenience sample within the school environment, which resulted in a male dominated sample consistent with the reported male domination of those who receive a diagnosis (for example, our ratio of 3:1 is close to Baird et al. (2006) who report 3.3:1). The matching of ASD and TD groups on cognitive ability in such a novel study was also not ideal as mathematical ability level was derived from teacher report and not formal independent testing (again within the classroom context, additional assessments were not feasible). With suboptimal matching we cannot discount the possibility that group differences in metacognitive ability were the result of general cognitive differences rather than diagnostic group membership per se. Independent assessments of both autism diagnosis and mathematical abilities would enable a more fine-grained analysis which would be useful additions to future research. However, when examining the relationship between metacognition and mathematics performance, the present study's focus upon being context-specific to the classroom may prove crucial.

To conclude, the present study reports a novel investigation of the impact of metacognitive support on autistic children's mathematics performance – the results of which have the potential to make a significant impact on autism education. It is now widely recognised that there is a vital need for evidence-based guidance to enable improvements in educational provision for pupils with ASD (e.g., Charman, et al., 2011; Keen et al., 2016; Wilkinson & Twist, 2010). An existing body of research indicates a metacognitive deficit in ASD, coupled with underachievement in mathematics that is in disaccord with intellectual ability. The present findings add to this and indicate that support for metacognitive assessments of accuracy and post-test intention monitoring may be particularly crucial targets for supporting learners with ASD in mathematics, with the potential to remove widely reported under achievement in mathematics.

References

- Aagten-Murphy, D., Attucci, C., Daniel, N., Klaric, E., Burr, D., & Pellicano, E. (2015). Numerical Estimation in Children With Autism. *Autism Research*, 8, 668–81. doi:10.1002/aur.1482
- Adcock, J., & Cuvo, A. J. (2009). Enhancing learning for children with autism spectrum disorders in regular education by instructional modifications. *Research in Autism Spectrum Disorders*, 3, 319–328. doi:10.1016/j.rasd.2008.07.004
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author
- Andersen, E. (2012). Optimizing adaptivity in educational games. In M. S. El-Nasr, M. Consalvo & S. Feiner (Eds.), *Proceedings of the International Conference on the Foundations of Digital Games* (pp. 279–281). New York, NY: ACM. doi:10.1145/2282338.2282398
- Azevedo, R., & Bernard, R. M. (1995). A Meta-Analysis of the Effects of Feedback in Computer-Based Instruction. *Journal of Educational Computing Research*, 13, 111–127. doi:10.2190/9LMD-3U28-3A0G-FTQT
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*, 368, 210–215. doi:10.1016/S0140-6736(06)69041-7.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61, 213–238. doi:10.3102/00346543061002213
- Barbeau, E. B., Soulières, I., Dawson, M., Zeffiro, T. A., & Mottron, L. (2013). The level and nature of autistic intelligence III: Inspection time. *Journal of Abnormal Psychology*, 122, 295–301. doi:10.1037/a0029984
- Bebko, J. M., & Ricciuti, C. (2000). Executive Functioning and Memory Strategy Use in Children with Autism: The Influence of Task Constraints on Spontaneous Rehearsal. *Autism*, 4, 299–320. doi:10.1177/1362361300004003006
- Boekaerts, M. & Corno, L. (2005) Self-Regulation in the Classroom: A Perspective on Assessment and Intervention. *Applied Psychology*, 54, 199–231. doi: 10.1111/j.1464-0597.2005.00205.x
- Brosnan, M., Johnson, H., Grawemeyer, B., Chapman, E., Antoniadou, K., & Hollinworth, M. (2015). Deficits in metacognitive monitoring in mathematics assessments in learners with autism spectrum disorder. *Autism* : doi:10.1177/1362361315589477
- Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, 46, 3–18. doi:10.1111/j.1469-7610.2004.00374.x
- Charlop, M. H., Kurtz, P. F., & Milstein, J. P. (1992). Too much reinforcement, too little behavior: assessing task interspersal procedures in conjunction with different reinforcement schedules with autistic children. *Journal of Applied Behavior Analysis*, 25, 795–808. doi:10.1901/jaba.1992.25-795
- Charman, T., Pellicano, L., Peacey, L. V., Peacey, N., Forward, K., & Dockrell, J. (2011). *Good Practice Report: What is Good Practice in Autism Education?* London: Autism Education Trust.
- Chiang, H.-M., & Lin, Y.-H. (2007). Mathematical ability of students with Asperger syndrome and high-functioning autism: a review of literature. *Autism*, 11, 547–56. doi:10.1177/1362361307083259
- Chong, I. M., & Carr, J. E. (2005). An investigation of the potentially adverse effects of task interspersal. *Behavioral Interventions*, 20, 285–300. doi:10.1002/bin.202
- Cornoldi, C., Lucangeli, D., Caponi, B., Falco, G., Focchiatti, R., & Todeschini, M. (1995). *Matematica e Metacognizione*. Trento: Erickson.
- Desoete, A., & Veenman, M. (2006). Metacognitions in mathematics: Critical issues on nature, theory, assessment and treatment. In A. Desoete & M. Veenman (Eds.), *Metacognition in mathematics education* (pp. 1–10). Hauppauge, NY: Nova Science.

- Dignath, C, Buettner, G and Langfeldt, H (2008). How can primary school students learn self-regulated learning strategies most effectively? A meta-analysis on self-regulation training programmes. *Educational Research Review*, 3, 101-129. doi:10.1016/j.edurev.2008.02.003
- Dunlap, G., & Koegel, R. L. (1980). Motivating autistic children through stimulus variation. *Journal of Applied Behavior Analysis*, 13, 619–27. doi:10.1901/jaba.1980.13-619
- Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging*, 18(2), 340-345.
- Efklides, A. (2011). Interactions of Metacognition With Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, 46, 6–25. doi:10.1080/00461520.2011.538645
- Estes, A., Rivera, V., Bryan, M., Cali, P., & Dawson, G. (2011). Discrepancies between academic achievement and intellectual ability in higher-functioning school-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(8), 1044-1052.
- Farrant, A., Blades, M., & Boucher, J. (1999). Recall readiness in children with autism. *Journal of Autism and Developmental Disorders*, 29(5), 359–66.
- Farrant, A., Boucher, J., & Blades, M. (1999). Metamemory in children with autism. *Child Development*, 70(1), 107–31.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906-911.
- Grainger, C., Williams, D.M. & Lind, S.E. (2016). Metacognitive monitoring and control processes in children with autism spectrum disorder: Diminished judgement of confidence accuracy. *Consciousness and Cognition*, 42, 65-74. doi:10.1016/j.concog.2016.03.003
- Grainger, C., Williams, D. M., & Lind, S. E. (2014). Metacognition, Metamemory, and Mindreading in High-Functioning Adults With Autism Spectrum Disorder. *Journal of Abnormal Psychology*, 123, 650–659. doi:10.1037/a0036531
- Griswold, D. E., Barnhill, G. P., Myles, B. S., Hagiwara, T., & Simpson, R. L. (2002). Asperger Syndrome and Academic Achievement. *Focus on Autism and Other Developmental Disabilities*, 17(2), 94–102. doi:10.1177/10883576020170020401
- Hartwig, M. K., Was, C. A., Isaacson, R. M., & Dunlosky, J. (2012). General knowledge monitoring as a predictor of in-class exam performance. *British Journal of Educational Psychology*, 82, 456–468.
- Higgins, S., Katsipatakis, M., Kokotsaki, D., Coleman, R., Major, L.E., & Coe, R. (2013). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit*. London: Education Endowment Foundation.
- Hunicke, R. (2005). The case for dynamic difficulty adjustment in games. In N. Lee (Ed.), *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (pp. 429-433). New York, NY: ACM. doi:10.1145/1178477.1178573
- Iuculano, T., Rosenberg-Lee, M., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., ... Menon, V. (2014). Brain organization underlying superior mathematical abilities in children with autism. *Biological Psychiatry*, 75, 223–30. doi:10.1016/j.biopsych.2013.06.018
- Isaacson, R.M. & Fujita, F. (2006). Metacognitive Knowledge Monitoring and Self-Regulated Learning: Academic Success and Reflections on Learning. *Journal of the Scholarship of Teaching and Learning*, 39 - 55. doi: 10.1.1.526.6161
- Jarrold, C. & Brock, J. (2004) To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, 34, 81-86. doi: 10.1023/B:JADD.0000018078.82542.ab
- Jones, C. R. G., Happé, F., Golden, H., Marsden, A. J. S., Tregay, J., Simonoff, E., ... Charman, T. (2009). Reading and arithmetic in adolescents with autism spectrum disorders: peaks and dips in attainment. *Neuropsychology*, 23, 718–28. doi:10.1037/a0016360

- Keen, D., Webster, A., & Ridley, G. (2016). How well are children with autism spectrum disorder doing academically at school? An overview of the literature. *Autism*, 20, 276–94. doi:10.1177/1362361315580962
- Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding executive control in autism spectrum disorders in the lab and in the real world. *Neuropsychology Review*, 18, 320–38. doi:10.1007/s11065-008-9077-7
- Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics intervention for children with special educational needs. *Remedial and Special Education*, 24, 97–114.
- Lockl, K., & Schneider, W. (2007). Knowledge about the mind: links between theory of mind and later metamemory. *Child Development*, 78, 148–67. doi:10.1111/j.1467-8624.2007.00990.x
- Maxwell, B. R., & Grenier, K. (2014). The Effects of Metacognitive Treatments on the Academic Performance of Students with Learning Disabilities: A Meta-Analysis. *Canadian Journal for New Scholars in Education/ Revue Canadienne Des Jeunes Chercheurs En Education*. Retrieved from <http://cjnse.journalhosting.ucalgary.ca/ojs2/index.php/cjnse/article/view/285>
- Mayes, S. D., & Calhoun, S. L. (2006). Frequency of reading, math, and writing disabilities in children with clinical disorders. *Learning and Individual Differences*, 16, 145–157. doi:10.1016/j.lindif.2005.07.004
- Mayes, S. D., & Calhoun, S. L. (2003). Analysis of WISC-III, Stanford-Binet:IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders*, 33, 329–41. doi:10.1023/A:1024462719081
- McMahon, C., Henderson, H., Newell, L., Jamie, M., & Mundy, P. (2016). Metacognitive awareness of facial affect in higher-functioning children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46, 882–98, doi: 10.1007/s10803-015-2630-3
- Metcalfe, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18, 159–163. doi:10.1111/j.1467-8721.2009.01628.x
- Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the “labor-in-vain effect”. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 14(4), 676–86.
- Nelson, T. O., and Narens, L. (1990). Metamemory: a theoretical framework and new findings. *Psychology of Learning and Motivation*, 26, 125–141. doi: 10.1016/S0079- 7421(08)60053-5
- Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology*, 16, 337–348. doi:10.1111/j.2044-835X.1998.tb00756.x
- Pintrich, P.R. & de Groot, E.V. (1990) Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33–40. doi:10.1037/0022-0663.82.1.33
- Roebbers, C. M., Krebs, S. S., & Roderer, T. (2014). Metacognitive monitoring and control in elementary school children: Their interrelations and their role for test performance. *Learning and Individual Differences*, 29, 141–149. doi:10.1016/j.lindif.2012.12.003
- Roebbers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, 7. doi: 10.1007/s11409-012-9089-9
- Schneider, W., & Artelt, C. (2010). Metacognition and mathematics education. *ZDM: The International Journal on Mathematics Education*, 42, 149–161. doi:10.1007/s11858-010-0240-2
- Schraw, G. & Dennison, R.S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475.

- Schraw, G. & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351-371.
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78, 153-189. doi:10.3102/0034654307313795
- Special Education Support Service (SESS). (2009). Metacognition for the classroom and beyond: Differentiation and support for learners. Retrieved from http://www.sess.ie/sites/all/modules/wysiwyg/tinymce/jscripts/tiny_mce/plugins/filemanager/files/Projects/Equality_of_Challenge/SESS_Metacognition_Resource_V1.pdf
- Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning*, 19(1), 46-55.
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multi-trial learning. *Psychonomic Bulletin Review*, 6, 662-667.
- Van der Stel, M., & Veenman, M. V. J. (2010). Development of metacognitive skillfulness: A longitudinal study. *Learning and Individual Differences*, 20(3), 220-224.
- Veenman, M. V., Kok, R., & Blöte, A. W. (2005). The relation between intellectual and metacognitive skills in early adolescence. *Instructional Science*, 33(3), 193-211.
- Vlamings, P. H. J. M., Jonkman, L. M., Hoeksma, M. R., van Engeland, H., & Kemner, C. (2008). Reduced error monitoring in children with autism spectrum disorder: an ERP study. *The European Journal of Neuroscience*, 28, 399-406. doi:10.1111/j.1460-9568.2008.06336.x
- White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research*, 2, 138-47. doi:10.1002/aur.78
- Whyte, E. M., Smyth, J. M., & Scherf, K. S. (2015). Designing Serious Game Interventions for Individuals with Autism. *Journal of Autism and Developmental Disorders*, 45(12), 3820-3831. <http://doi.org/10.1007/s10803-014-2333-1>
- Wilkinson, K. and Twist, L. (2010). *Autism and Educational Assessment: UK Policy and Practice*. Slough: NFER.
- Williams, D. M., Bergström, Z., & Grainger, C. (2016). Metacognitive monitoring and the hypercorrection effect in autism and the general population: Relation to autism(-like) traits and mindreading. *Autism*. doi:10.1177/1362361316680178
- Williams, D., & Happé, F. (2010). Representing intentions in self and other: studies of autism and typical development. *Developmental Science*, 13, 307-19. doi:10.1111/j.1467-7687.2009.00885.x
- Wojcik, D. Z., Moulin, C. J. A., & Souchay, C. (2013). Metamemory in children with autism: exploring "feeling-of-knowing" in episodic and semantic memory. *Neuropsychology*, 27, 19-27. doi:10.1037/a0030526
- World Health Organization. (1993). *The international statistical classification of diseases and related mental health problems, tenth revision (ICD-10)*. Geneva: World Health Organization
- Xojo. (2011). Real Studio (Version 4.3) [Computer software]. Austin, TX: Author.
- Zimmerman, B. J., & Tsikalas, K. E. (2005). Can computer-based learning environments (CBLEs) be used as self-regulatory tools to enhance learning? *Educational Psychologist*, 40(4), 267-271.